EVOLUTIONARY OPTIMIZATION OF GEOMETRIC TOLERANCES FOR COMPLIANT ASSEMBLIES WITH CONTACT INTERACTIONS

Arunachalam Saravanan¹, C. Balamurugan¹ and A. Karthik²
¹Department of Mechanical Engineering, M.A.M College of Engineering, Trichy, Tamilnadu, India
²PRIST University, Tanjore, Tamilnadu, India
Email: a_saravanan@mamce.org

ABSTRACT

Geometric Dimensioning and Tolerancing (GD and T) is an important activity during product development phase, since it directly influences the manufacturing cost and subsequent manufacturing time. Traditional GD and T design methodology assumes that the complete assembly is perfectly rigid; hence it often leads to costly reworks and rejections. This increases the product development time and cost, also it forces the entire assembly to accommodate major Engineering Changes (EC) in order to meet the product functionality. This paper proposes an ideal alternative methodology to overcome the effects of the assumptions. Initially contact interactions in the assembly are simulated by Finite Element Analysis (FEA), since the mechanical contact influences severe nonlinearities, which makes the assembly compliant and non-ideal. The simulated compliant assembly is modeled using three dimensional degrees of freedom approach and an optimization problem is framed with an objective function of minimizing the manufacturing cost. Assembly function constraints, machining constraints and tolerance zone constraints are considered to solve the problem using Genetic Algorithm (GA). Finally an industrial example is chosen to illustrate the highly structured methodology.

Keywords: compliant assembly, engineering tolerance optimization, finite element analysis (FEA), genetic algorithm (GA), geometric dimensioning and tolerancing (GD and T).

INTRODUCTION

Geometric Dimensioning and Tolerancing (GD and T) is a precise mathematical language that can be used to specify the acceptable variation for a part/assembly [1]. GD and T is a set of symbols/rules with multi functional aspect that correlates varies activities of a product at design stage; the main aim of using GD and T is to minimize manufacturing cost and time with enhanced bonus tolerances. Traditional GD and T design methodology was used in detailed drawings to reduce the manufacturing cost of that particular part [2-4]. Later researchers focused on the application of GD and T concepts during product development phase and consequent system level design [5, 6, 7]. When GD and T is used for assemblies there raised conflict between design and manufacturing. Designers preferred close tolerance considering fit, function and performance while manufacturing engineers expected liberal tolerances to meet their production schedule. Hence tradeoff between design and manufacturing became essential and several researchers optimized the scenario using various algorithms. [8-10]. Thus the design methodology was scrutinized by researchers with following assumptions [11]. 1. Each process follows normal distribution and is under statistical control. 2. The dimensions in a dimension chain and the process for each dimension are independent. 3. Geometric tolerance for a feature is considered for its Actual mate envelope (AME) and corresponding tolerance zone, not for the design feature. 4. Material condition is assumed as AME and feature is obtained directly from machine. 5. The components in the assembly are perfectly rigid and are free from deformation/variations.

The assumptions are good for an individual component and are manufactured with least possible cost. When the parts are assembled they completely deviate from the functional requirements and are sub standard. Thus costly rejections or rework need to be executed for functionality. Such forced modifications are called Engineering Changes (EC) and becomes mandate for making the assembly right. The root cause for such EC is the assumptions considered in the assembly based GD and T design; that the components in the assembly are perfectly rigid. This paper proposes an ideal alternative methodology to overcome the ill effects of the assumption and optimize the tolerance values with a Functional Assembly (FA).

PROPOSED ASSEMBLY DESIGN PROCEDURE

Figure-1. Simplified tolerance synthesis model [12].
This section details the proposed procedure to address the GD and T related issues during assembly and associated assumptions. The main scope of the paper is to break the assumptions assumed during the application of GD and T in assembly design. When a component is assembled, it is always subjected to certain force and it’s a non-ideal called compliant component. As a result it experiences distortion, the role of distortion need to be considered while allotting tolerances for that component [12]. Figure-1 shows the new procedure modeled by Manarvi and Juster for including the role of distortion in tolerance synthesis. The effect of distortion was calculated through Finite Element (FE) simulation using software visual tools. Thus FE simulation helps us to overcome the assumption of rigid parts. On accounting the sources of variation during assembly are Inertia (gravity, velocity), Temperature, external forces, self weight or combination of the above [13]. The complete procedure for FE based tolerance design has been demonstrated with non linear GD and T optimization has been explained in next section.

STEPS FOR GD AND T ASSEMBLY DESIGN PLATFORM

The typical GD and T assembly design platform approach involves the following 10 steps.
1. Definition of CAD assembly model with individual constraint parts.
2. Specification of nonlinear contact interactions in the assembly.
3. Definition of part relations such as assembly sequence, mating conditions etc.
4. Importing of the neutral file to a FE package.
5. FE Simulation of the compliance element with appropriate loads and constraints.
6. Recording of the outcomes and analysis of the data.
7. Obtaining the functional model in the FE package database.
8. Establishing the objective function and relevant constraints.
9. Running the Genetic Algorithm (GA) with the data described in step-8.
10. Analyze and implement the results in the CAD model.

The above 10 steps are distinctly grouped under 2 phases. First sets up the FE simulation activity and second involves modeling and optimization.

FINITE ELEMENT SIMULATION

The first phase considers the variations in a system arising from different sources like design, part and subsequent assembly (Figure-2). To demonstrate the functioning of the proposed method a case study is shown in Figure-3 [13]. It is vane driven pump assembly used in hydraulic assisted power steering mechanism. Farmer et al claims that the assembly focuses severe EC at all levels of processing. Rework cost equally affecting the manufacturing cost also quality is comprised.
MESHING

Meshing involves approximating the actual physical structure using several simple geometric shapes called elements. These elements are interconnected to each other at points called Nodes. The mesh is a representation of the mathematical idealization of the structure. Within each element, displacement of nodes is determined by a polynomial equation called displacement equation. For this problem a higher order OCTREE Tetrahedron mesh was defined for all the parts except nonlinear interacting areas. Parabolic Tetrahedron elements (PTE) shown in Figure-4 were used in interaction areas; it is done to differentiate them as different nonlinear issues.

Figure-4. Parabolic tetrahedron element.

Tetrahedron elements use displacement interpolation such as parabolic, cubic or higher order between the nodes. Thus when parabolic elements are subjected to loads their shapes follow parabolic deformation equation, these elements have additional nodes on the edges joining the primary nodes. They are used to improve the accuracy of the solution; however they increase the computation time. The accuracy of the result is primarily affected by the mesh quality. The initial PTE mesh size was not uniform as shown in Figure-5, also sag dimension was also to large hence the mesh was refined with uniform mesh size (5mm) and Small sag (0.026 mm) as shown in Figure-6 with 3D property. The refined high quality FE model is subjected to restraints and loads.

Figure-5. PTE meshed assembly.

RESTRAINTS

Restraints are used to specify the support or boundary conditions for the FE model; they restrict the displacement of supports of a structure in the desired direction. This is done by providing zero displacement values for specific Degrees of Freedom (DOF) of the nodes in FE model. The restraints are directly applied onto the geometry. The meshed assembly is defined with Isostatic restraint, which represents a simply supported body. The resulting boundary condition prevents the body from rigid-body translations and rotations without over constraining it.

LOADS

Loads are main inputs to the FE model. There are various types of loads available. In this case study distributed forces have been implied. These forces are statically equivalent to a given pure force at a given point, distributed on a virtual part or on a geometric selection. Figure-7 shows the restrained and loaded assembly.

Before displacements matrices. This case study used Gauss R6 method for computation. The obtained results have been
checked which allows verifying whether all the pre-processing steps are done and if the model is ready for computation.

**COMPUTATION**

Computation is required to calculate the unknown displacement values at the nodal points of the FE model. From these displacement values other solution quantities such as strain, stresses are derived. Initially the geometry is discretized into elements; all properties and forces are idealized at the element and nodal level. For each element, nodal forces, stiffness matrices and unknown displacement vectors are computed. The element connectivity is used to assemble the global stiffness, nodal forces and displacements matrices. This case study used Gauss R6 method for computation. The obtained results have been shown in Figures 8, 9 and tabulated Table-2. The deformed zone is the deviation from the basic size of the shaft from x_t model. From the deformed nodal points of the assembly a CATIA graphic report (*.CGR) file is created. These nodal points are used as coordinates and points are plotted to form the outer topography and then the envelope. The envelope outer dimensions are measured and difference between the initial geometry is taken as the envelope/deformed zone dimensions.

![Figure-8. Deformed nodes of the model.](image)

![Figure-9. Deformed non linear shafts.](image)

**Table-2. Result summary.**

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Part</th>
<th>Deformed Zone</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Small shaft</td>
<td>0.18mm x 0.8mm</td>
</tr>
<tr>
<td>2</td>
<td>Longer shaft</td>
<td>0.225mm x 1.2mm</td>
</tr>
</tbody>
</table>

**NON-LINEAR OPTIMIZATION**

For optimal determination of geometric tolerances, an optimization problem with appropriate tradeoff between assembly function and manufacturing cost is formulated as follows.

**OBJECTIVE FUNCTION**

The objective function minimizes the total cost in an assembly. Here a cost-process tolerance function is adopted as the manufacturing cost component of the objective function. Each manufacturing operation is modeled with an appropriate geometric tolerance relationship. This avoids the inaccuracies of cost-design tolerance models, and permits direct distribution of design tolerances to each process tolerance. The total cost is the sum of manufacturing cost of each component’s tolerance.

\[
\text{Minimize} \quad \sum_{j=1}^{n} \sum_{i=1}^{d} C_j(t_i) \quad (1)
\]

Where

\[
C_j(t_i) = \text{manufacturing cost of tolerance } t_i \text{ for model } j \\
t_i = \text{tolerance value of the features.}
\]

**MANUFACTURING COST-TOLERANCE FUNCTION**

It describes the manufacturing cost incurred to produce the assigned geometrical tolerance. Several formulations of machining cost-tolerance models for modeling the cost–tolerance relationship such as exponential model, inverse square model, inverse power model, and inverse model have been developed and reviewed by different researchers [14]. Although non-traditional cost functions model the characteristics of the manufacturing processes more accurately, for a balance between modeling accuracy and computational simplicity, the exponential cost function model used by Hu and Xiong [15] is considered the best and the machining cost-tolerance relation is broadly classified into 4 models.

**Model-1:** Manufacturing cost-tolerance function for size tolerances \((\pm)\) for shaft feature (external cylinder)

\[
C_1(t_i) = 10^{-5} + 10^{-7}t_i + 67.3e^{-2.559t_i} \quad (2)
\]

**Model-2:** Manufacturing cost-tolerance function for size tolerances \((\pm)\) for hole feature (internal cylinder)

\[
C_2(t_i) = 10^{-5} + 10^{-7}t_i + 57.6e^{-1.559t_i} \quad (3)
\]
Model-3: Manufacturing cost-tolerance function for positional tolerance for cylindrical feature

\[ C_3(t_i) = 0.032 + 10^{-5} t_i + 30.07 e^{-12.06 t_i} \]  

(4)

Model-4: Manufacturing cost-tolerance function for perpendicular tolerance

\[ C_4(t_i) = 5.425 + 10^2 t_i + 12.43 e^{10.06 t_i} \]  

(5)

CONSTRANTS

The above-mentioned non-linear optimization problem is subjected to four constraints related to both the design and manufacturing attributes. They are:

1. Fit function constraint,
2. Cylindrical machining constraints,
3. Surface machining constraints and
4. Tolerance zone constraints.

The above constraints were modeled using Kinematics principle and detailed in following section.

CONSTRAINT MODELING

There are 4 major constraints formulated for the above optimization problem, they are stack-up constraints, translational constraints and tolerance zone associated with the design constraints. The preliminary phase of understanding these 3D constraints for GD and T design is by interpreting the Degrees of freedom (DOF) available for a part when kept in space as illustrated in Figure-10. There are 6 DOFs available for a part when kept in space. They are 3 independent translations in x, y and z axis Translation movements (T_x, T_y, T_z), clockwise rotations (R_x, R_y, R_z) and counter clockwise rotations (R_x, R_y, R_z) = key parameters (u_i, v_i, w_i) for translation along x, y, z axis Rotational movements (R_x, R_y, R_z) = key parameters (a_i, b_i, c_i) for rotation along x, y, z axis. The problem with respect to tolerance design the values in each axis will be very small so they are designated by key parameters. Translational movements (T_x, T_y, T_z) = key parameters (u_i, v_i, w_i) for translation along x, y, z axis Rotational movements (R_x, R_y, R_z) = key parameters (a_i, b_i, c_i) for rotation along x, y, z axis. Key matrix is the tolerance constraints that can be formed using these parameters and it is defined as:

\[ \mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix} \]  

(6)

where

\[ \mathbf{t} = \begin{bmatrix} a_i \\ b_i \\ c_i \end{bmatrix} \]  

is rotational key vector and

\[ \mathbf{t} = \begin{bmatrix} u_i \\ v_i \\ w_i \end{bmatrix} \]  

is directional or translational key vector are established with 3D parameters with respect to rotational and translational of a part in space.

FIT FUNCTION CONSTRAINT (FFC)

From this study it is clear that stack-up is classified into rotational stack-up and translational stack-up, will be discussed.

On considering two parts kept above each other in Figure-10 then the resultant Key matrix for the assembly is:

\[ \mathbf{t}_{\text{assembly}} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix} \]  

(7)

From the key matrix, it is possible to obtain the AFR constraint equation based Root Sum Square (RSS) approach. Krulikowski [16] presented RSS approach in his book as an approach is employed to account for the low likelihood of all dimensions occurring at their extreme limits simultaneously. The sum of squares is a mathematical treatment of the data to facilitate the legitimate addition of measures of variability. The RSS method is used to determine if a functional fit is going to occur between the mating assemblies. It is assumed that the sample data we are working which comes from reasonable approximations of normal distribution. With respect to RSS approach the FFC constraint equation is

\[ (v_i + L_i - \alpha_i)^2 + (u_i + L_i - \beta_i)^2 \leq (t_i)^2 \]  

(8)

Where

- \( L_i \) = feature of size (FOS) related to AFR in x axis,
- \( \alpha_i \) = tolerance value fixed by the designer for AFR
- \( \beta_i \) = key parameters assigned a value in the range of 0.1mm to 0.001mm.

CYLINDRICAL MACHINING CONSTRAINTS (CMC)

Using the rotational key parameters (\( \alpha_i, \beta_i, \gamma_i \)) along x, y, z axis the CMC equations are formulated to respective axis.

\[ \alpha_i = \sum_{i=1}^{n} \alpha_{i} \]  

(9)

\( \alpha_0 \) is the cumulative rotational stack-up for \( \alpha_i \) (i = 1, 2, 3 . . . n) along x axis.
$\beta_0$ is the cumulative rotational stack-up for $\beta_i$ ($i = 1, 2, 3 \ldots n$) along y axis.

$\gamma_0$ is the cumulative rotational stack-up for $\gamma_i$ ($i = 1, 2, 3 \ldots n$) along z axis.

SURFACE MACHINING CONSTRAINTS (SMC)

Translational stack-up is a state of constraint for an assembly at which minute and cumulative build up of deviation along x, y, z axis machining. Hu and Xiong [15] proposed the translational stack-up as:

$$u_0 = \sum_{i=1}^{n} u_i - \sum_{i=1}^{n} \gamma_i + \sum_{i=1}^{n} \beta_i Z_i$$

$v_0$ is the translational constraint along y axis.

$$w_0 = \sum_{i=1}^{n} w_i - \sum_{i=1}^{n} \gamma_i + \sum_{i=1}^{n} \alpha_i Z_i$$

$w_0$ is the translational constraint along z axis.

TOLERANCE ZONE CONSTRAINTS (TZC)

The part which is shown in Figure-10 is a design; there exist certain deviation when observed practically. Srinivasan [17] studied this tolerance zone theory and presented a geometrical product specification language for computer-aided tolerancing. The representation used in the algorithm is based on the study of variational model using key matrix (Figure-11).

OPTIMIZATION PROBLEM

This section describes the final optimization problem considering the models discussed in above section.

OBJECTIVE FUNCTION

Mathematical expression is essential for minimizing the relative manufacturing cost to produce that tolerance. Hence ‘i’ ranges from 1 to 10 (Since 10 tolerances) and expressed as:

$$c_{j_{ti}} = c_1(t_1) + c_1(t_2) + c_1(t_3) + c_4(t_5) + c_4(t_6) + c_2(t_7) + c_2(t_8) + c_2(t_9) + c_2(t_{10})$$

The objective function is made as a combinatorial function by including the relevant ‘c’ models as obtained from equation (2) to (5). Hence combinatorial equations of this objective function is:

$$\text{Minimize: } \left[ \sum_{j=1}^{10} \sum_{t=1}^{4} c_{j_{ti}} \right]$$

FIT FUNCTION CONSTRAINT (FFC)

$$\begin{align*}
(\gamma_0 + L_{\beta_0})^2 + (u_0 - L_{\beta_0})^2 & \leq (AFF)^2 \\
(\gamma_0 + 25 \alpha_0)^2 + (u_0 + 25 \alpha_0)^2 & \leq (0.020)^2 \\
g_0 - (0.020)^2 & \geq 0
\end{align*}$$
Dynamic optimization is time dependent, whereas the process of optimization becomes difficult. As the number of dimensions increases, the process of optimization becomes difficult. A mathematical formula describes the objective function for optimization. One dimensional optimization contains one variable and a problem having more than one variable requires multi-dimensional approach. As the number of dimensions increases, the process of optimization becomes difficult. Dynamic optimization is time dependent, whereas the optimization algorithms can be characterized into five categories. In trial and error optimization, the processes needed to tackle the unsolvable or hard problems. Genetic algorithms are computer based search techniques patterned after the genetic mechanisms of biological organisms that have adapted and flourished in changing highly competitive environment. Last decade has witnessed many exciting advances in the use of genetic algorithms to solve optimization problems in process control systems. Genetic algorithms are the solution for optimization of hard problems quickly, reliably and accurately. As the complexity of the real-time controller increases, the genetic algorithms applications have grown in more than equal measure. One of the most fundamental principal in our world is the search for an optimal state. Optimization is the process of modifying the inputs or characteristics of a device, mathematical process to obtain minimum or maximum of the output. The input to the optimization process is the cost function, objective function or fitness function and the output is the fitness function of the system. Optimization is a primary tool, responsible to produce the output. A mathematical formula describes the objective function for optimization. One dimensional optimization contains one variable and a problem having more than one variable requires multi-dimensional approach. As the number of dimensions increases, the process of optimization becomes difficult. Dynamic optimization is time dependent, whereas the static optimization is independent of time. The static problem is difficult to solve for finding the best solution but the added dimension of time increases the challenge of solving dynamic problems. Discrete variable optimization contains only a finite number of possible values, whereas continuous variables have an infinite number of possible values. Variables often have limits or constraints. Constrained optimization incorporates variable equalities and inequalities into the cost function, whereas unconstrained optimization allows the variable to take any value. A constrained optimization problem can be converted into unconstrained one through the transformation of variables. Many optimization algorithms have been developed in their original form. The goal of global optimization is to find the global optima, that is, global maxima or minima of the objective function. Optimization problems are used to find good parameters or designs for components or plans to be put into action by the human beings or machines.

WORKING PRINCIPLE OF GA
The workability of GA is based on Darwinian’s theory of survival of the fittest [19 and 20]. GA may contain a chromosome, a gene, set of population, fitness, fitness function, breeding, mutation and selection. GA begins with a set of solutions represented by chromosomes, called population. Solutions from one population are taken and used to form a new population, which is motivated by the possibility that the new population will be better than the old one. Further, solutions are selected according to their fitness to form new solutions, that is, offsprings. The above process is repeated until some condition is satisfied. Algorithmically, the basic GA is outlined as below:

**Step-I:** [Start] Generate random population of chromosomes, that is, suitable solutions for the problem.

**Step-II:** [Fitness] Evaluate the fitness of each chromosome in the population.

**Step-III:** [New population] Create a new population by repeating following steps until the new population is complete.

a) [Selection] Select two parent chromosomes from a population according to their fitness. Better the fitness, the bigger chance to be selected to be the parent.

b) [Crossover] with a crossover probability, cross over the parents to form new offspring, that is, children. If no crossover was performed, offspring is the exact copy of parents.

c) [Mutation] with a mutation probability, mutate new offspring at each locus.

d) [Accepting] Place new offspring in the new population.

**Step-IV:** [Replace] Use new generated population for a further run of the algorithm.

**Step-V:** [Test] If the end condition is satisfied, stop, and return the best solution in current population.

**Step-VI:** [Loop] Go to step 2.
The genetic algorithms performance is largely influenced by crossover and mutation operators. The flow chart representation of genetic algorithms (GA) is shown in Figure-12.

![Figure-12. GA working flow chart.](image)

**PARAMETERS USED IN GA**

The problem was solved using the following parameters:

1. Population size: 100
2. No. of iteration: 100
3. Cross-over probability: 0.7
4. Mutation probability: 0.6
5. Distribution index for cross over: 10
6. Distribution index for mutation: 100

![Figure-13. Iteration in GA.](image)

Table-3. The manufacturing reference cost from the algorithm is 378 Cr. These values are determined over the envelope size and therefore the functionality is maintained through the product life and forced EC is not required anymore. There by the proposed approach eliminates the costly reworks and rejections.

<table>
<thead>
<tr>
<th>S. No.</th>
<th>Tolerance code</th>
<th>Tolerance value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>T₁</td>
<td>0.0195</td>
</tr>
<tr>
<td>2</td>
<td>T₂</td>
<td>0.0188</td>
</tr>
<tr>
<td>3</td>
<td>T₃</td>
<td>0.0199</td>
</tr>
<tr>
<td>4</td>
<td>T₄</td>
<td>0.0189</td>
</tr>
<tr>
<td>5</td>
<td>T₅</td>
<td>0.0200</td>
</tr>
<tr>
<td>6</td>
<td>T₆</td>
<td>0.0150</td>
</tr>
<tr>
<td>7</td>
<td>T₇</td>
<td>0.0160</td>
</tr>
<tr>
<td>8</td>
<td>T₈</td>
<td>0.0200</td>
</tr>
<tr>
<td>9</td>
<td>T₉</td>
<td>0.0178</td>
</tr>
<tr>
<td>10</td>
<td>T₁₀</td>
<td>0.0166</td>
</tr>
</tbody>
</table>

**RESULTS**

The solution for the assembly considering the compliance and non linear interaction is tabulated in Table-3. The manufacturing reference cost from the algorithm is 378 Cr. These values are determined over the envelope size and therefore the functionality is maintained through the product life and forced EC is not required anymore. There by the proposed approach eliminates the costly reworks and rejections.

**CONCLUSION AND FUTURE SCOPE**

Traditionally, tolerance allocation is done based on a hypothesis that the assembly process deals with infinitely rigid bodies. The assembly functions are developed on this hypothesis. The resultant tolerance of individual components obtained will be on the tighter side, thereby increasing the manufacturing cost. In reality, all the components of the assembly are deformable bodies and undergo deformation due to distributed load. Through finite element simulation, the value of deformation has been determined. The deformation value is suitably incorporated in the constraint equation of the tolerance design problem. With the presented approach, the component tolerance values found are the most robust to variation of operating conditions during the product’s application. Due to this, the tolerance requirements of the given assembly are relaxed to certain extent for critical components, resulting in reduced manufacturing cost and high product reliability. These benefits make it possible to create a high-quality and cost-effective tolerance design, commencing at the earliest stages of product development. This problem has addressed the functional assembly considering the distributed load and appropriate restraints. Practically combination of external forces act like force and temperature, inertia and temperature etc., hence coupled field analysis is required. After analysis the nodal points are generated into a parasolid model through CGR format. This can be extended to CAT part, to enhance the study of complete topology and integrating the same with CMM for advanced product development. The application of other EAs like Differential Evolution (DE), Non-dominated sorting and GA - II (NSGA - II) may also be chosen for investigation.
REFERENCES


